

Music Trends Through the Generations

Code Crusaders

Dayna Cole

Joshua Covarrubias

Rachel Taylor

Iryna Vitkovych

Danielle Woodworth

Background

Music has been around in various forms for many many years, with some forms even dating back to over 35,000 years ago. Throughout time, music has been used within religions, to celebrate, to share messages, to inspire, to express oneself, and to generally entertain. Across the decades, how music is created, shared, consumed, and utilized has evolved rapidly and innovatively.

Since the start of the new millenia, music consumption has been digitized and consumers are now more likely to consume music via streaming than in previous generations' where consumption was via physical or analog mediums.

Depending on what (recent) generation you may have been born in, the consumption, listening habits, and popularity of genres or music types may differ. Our team was interested in analyzing music trends of the past to determine if its possible to predict the trends of the next generation to come.



Objectives

In order to analyze streaming data to better understand music trends across listening generations*, we will utilize data from one of the top music streaming companies in the world, [Spotify](#). This data was retrieved via [Spotify's API](#) as well as a [Kaggle dataset of Spotify songs](#) with different genres and their audio [features](#)**.

Once our data has been cleaned and merged, we were able to do initial data exploration based on our objectives. Our objectives aimed to lead us to the overarching question of “***Based on listening trends of the past, is it possible to determine what will be popular in the next generation (Generation Beta)***”

Generations* (Based on one's birth year)



Baby Boomers (1946-1964)



Generation X (1965-1979)



Millennials (1980-1994)



Generation Z (1995-2009)



Generation Alpha (2010 -2024)

Code Crusaders' Objectives

1. How has the number of tracks (songs) changed over time, specifically across each generation.
2. Understand how have features of tracks evolved over time? What features seem to be the most common in each listening generation.
3. Determine how genres have evolved and if there are consistent genres that stand the test of time.

Overview of Our Dataset

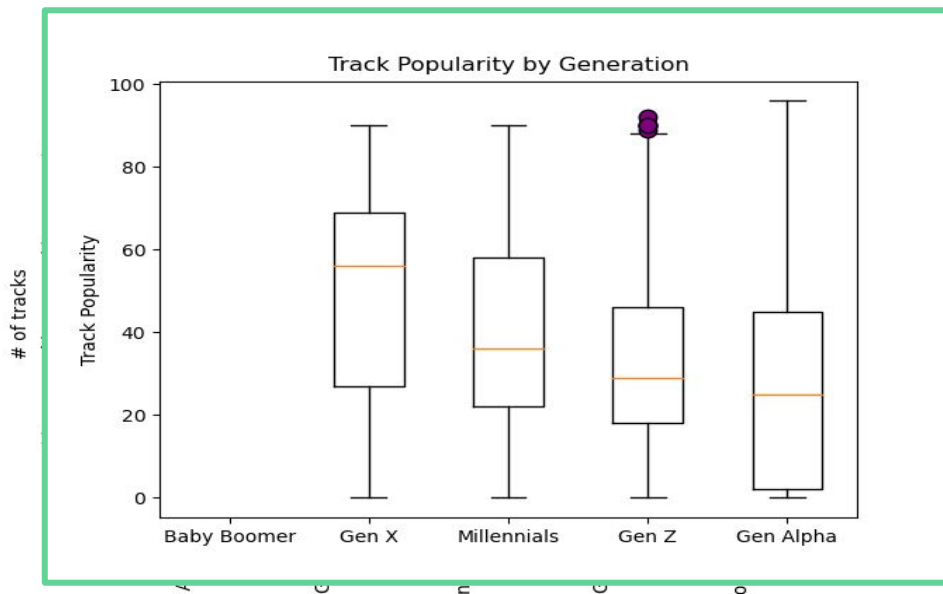
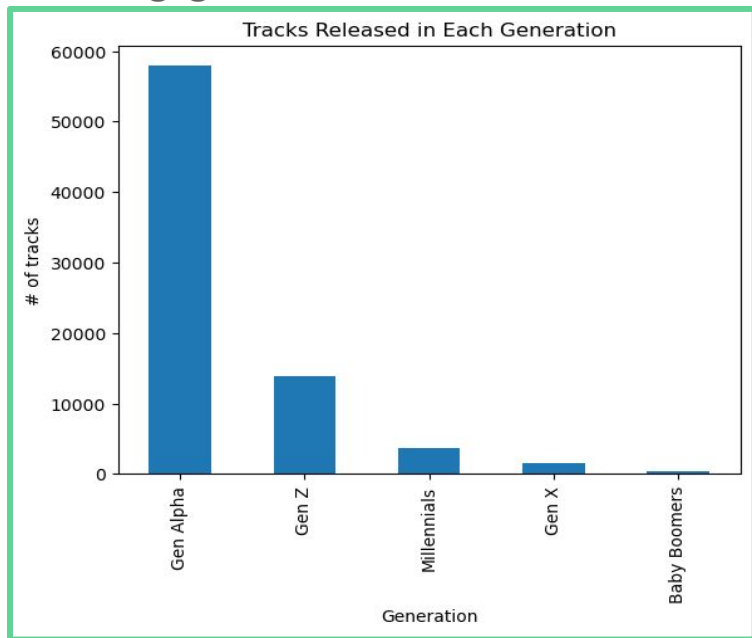
	track_id	artists	album_name	track_name	duration_sec	explicit	danceability	energy	key	loudness	...
0	5SuOikwiRyPMVoIQDJUgSV	Gen Hoshino	Comedy	Comedy	230.666	False	0.676	0.4610	1	-6.746	...
1	4qPNDBW1i3p13qLCt0Ki3A	Ben Woodward	Ghost (Acoustic)	Ghost - Acoustic	149.610	False	0.420	0.1660	1	-17.235	...
2	1iJBSr7s7jYXzM8EGcbK5b	Ingrid Michaelson;ZAYN	To Begin Again	To Begin Again	210.826	False	0.438	0.3590	0	-9.734	...
3	6lfxq3CG4xtTiEg7opyCyx	Kina Grannis	Crazy Rich Asians (Original Motion Picture Sou...	Can't Help Falling In Love	201.933	False	0.266	0.0596	0	-18.515	...
4	5vjLSffmilP26QG5WcN2K	Chord Overstreet	Hold On	Hold On	198.853	False	0.618	0.4430	2	-9.681	...
liveness	valence	tempo	time_signature	track_genre	release_date	#_of_tracks_in_album	track_popularity	release_year	artist		
0.3580	0.715	87.917	4	acoustic	2022-04-08	1.0	69.0	2022	Gen Hoshino		
0.1010	0.267	77.489	4	acoustic	2021-04-30	1.0	49.0	2021	Ben Woodward		
0.1170	0.120	76.332	4	acoustic	2021-03-17	1.0	54.0	2021	Ingrid Michaelson		
0.1320	0.143	181.740	3	acoustic	2018-08-10	14.0	70.0	2018	Kina Grannis		
0.0829	0.167	119.949	4	acoustic	2017-02-03	1.0	82.0	2017	Chord Overstreet		

Overview of Our Dataset Cont'd

1. Tracks were released anywhere between 1899-2023
 - a. Dataset was created in October of 2023 so we needed to grab current popularity scores
 - b. 114,000 rows of data
2. Created new columns
 - a. Release_date, release_year, #_of_tracks_in_album, and track_popularity
3. Spotify's API has limitations on the amount of data that you use.
 - a. 1 hour time limits on the API key
 - b. 30 second window call limit intervals
 - c. With the time limits to Spotify's API, we could only get data for 77,701 rows
4. Removed duplicate albums and song names
 - a. Includes any cover songs or live/extended/remastered versions.
5. Removed album or song names with any foreign letters and languages
 - a. With both methods of removing rows, this only got us down to around 81,827 rows
6. Lastly, removed 4,126 random rows

Initial Data Analysis Findings

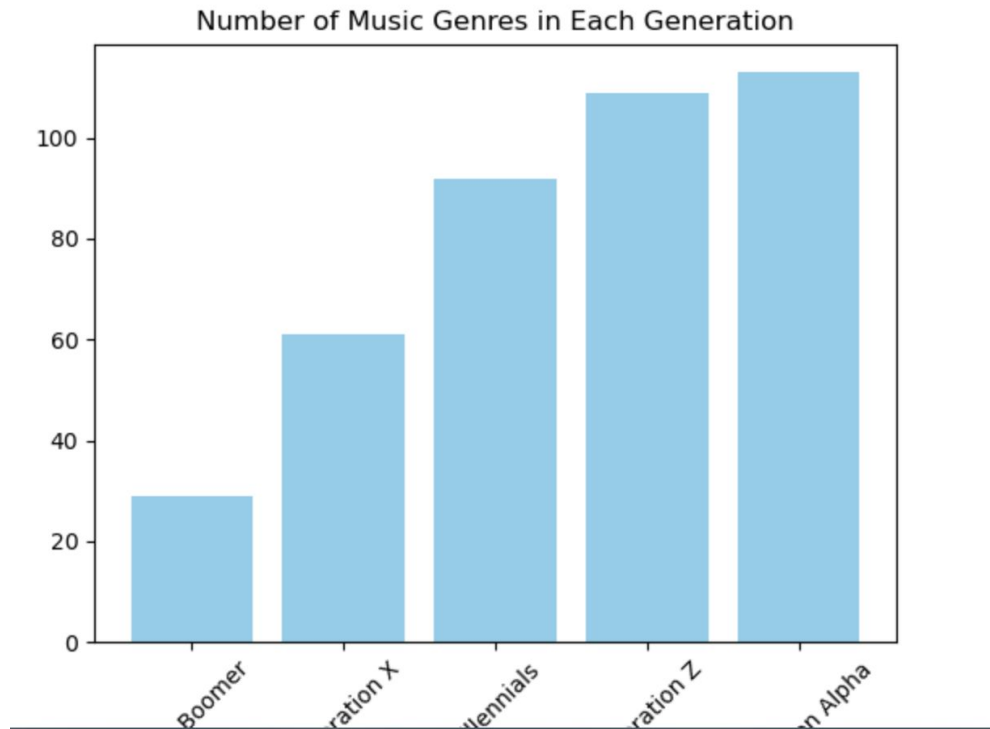
Across generations we started looking at the volume of tracks released in each listening generation.



Initial Data Analysis Findings Cont'd

We then look across generations to see the top genres for tracks released during each listening decade.

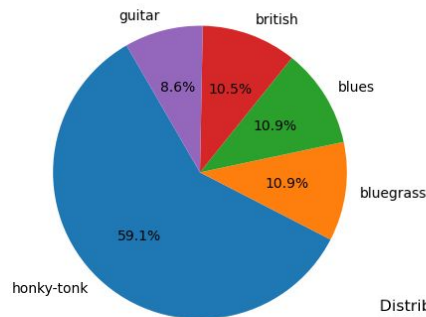
As we are currently in Generation Alpha, this year being Q1 2024, we can see that already the # of genres have increased since last period of Generation Z.



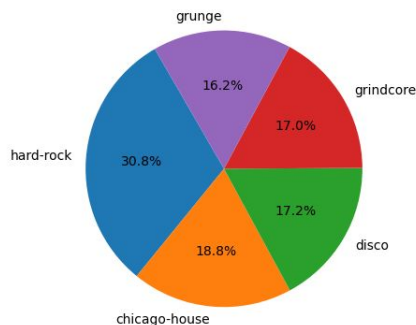
Initial Data Analysis Findings Cont'd

We then look across generations to see the top genres for tracks released during each listening decade

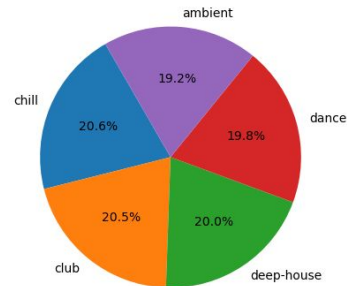
Distribution of Genres in the years of Baby Boomers



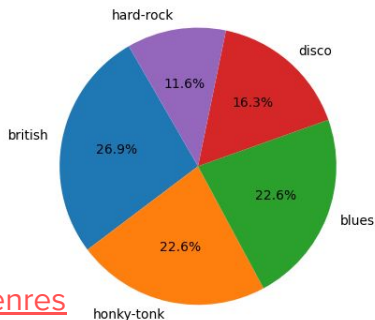
Distribution of Genres in the years of Millennials



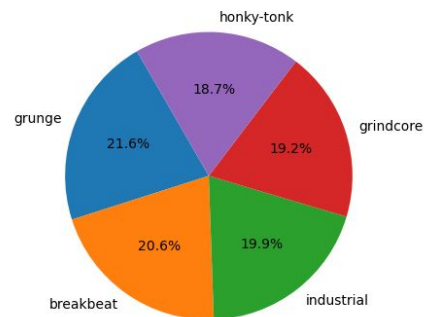
Distribution of Genres in the years of Gen Alpha



Distribution of Genres in the years of Gen X



Distribution of Genres in the years of Gen Z



Artists with the Most Top Tracks in a Generation

Baby Boomers (64 artists):

Artist Name	Number of Top Songs
Frank Sinatra	6
Elvis Presley	6
Ritchie Valens	4
Billie Holiday	4
Ray Charles	4

Millennials (101 artists):

Artist Name	Number of Top Songs
Guns N' Roses	6
Bon Jovi	4
Nirvana	4
R.E.M.	3
Fleetwood Mac	3

Gen Z (125 artists):

Artist Name	Number of Top Songs
Eminem	7
Rihanna	6
USHER	6
Coldplay	6
Kanye West	5

Gen X (108 artists):

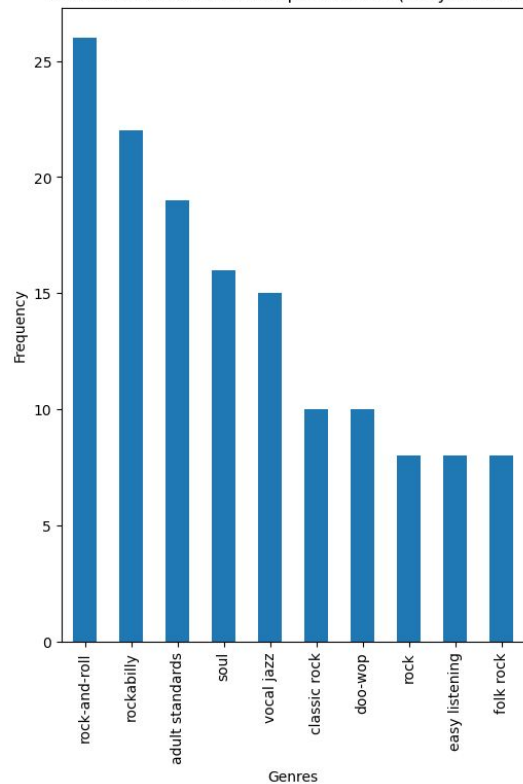
Artist Name	Number of Top Songs
ABBA	7
The Rolling Stones	7
Creedence Clearwater Revival	6
Led Zeppelin	5
Elton John	5

Gen Alpha (53 artists):

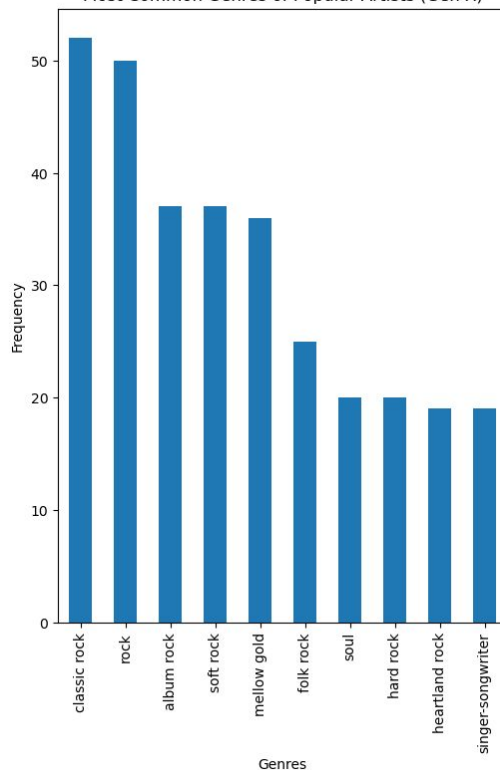
Artist Name	Number of Top Songs
The Weeknd	4
Ed Sheeran	4
Sam Smith	3
Kendrick Lamar	3
Calvin Harris	3

Top Artist Genres

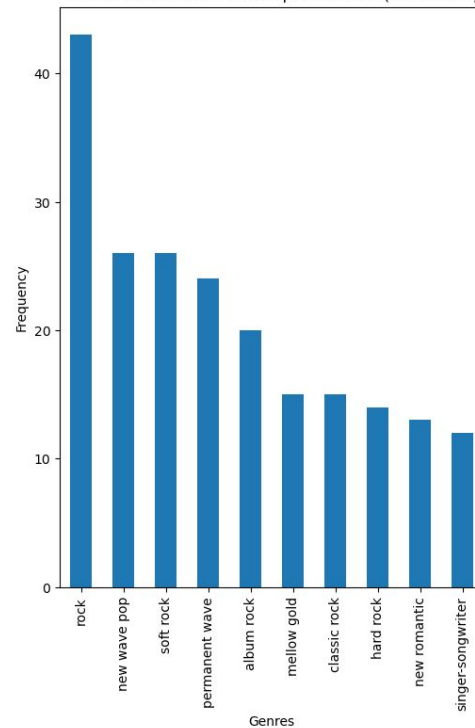
Most Common Genres of Popular Artists (Baby Boomers)



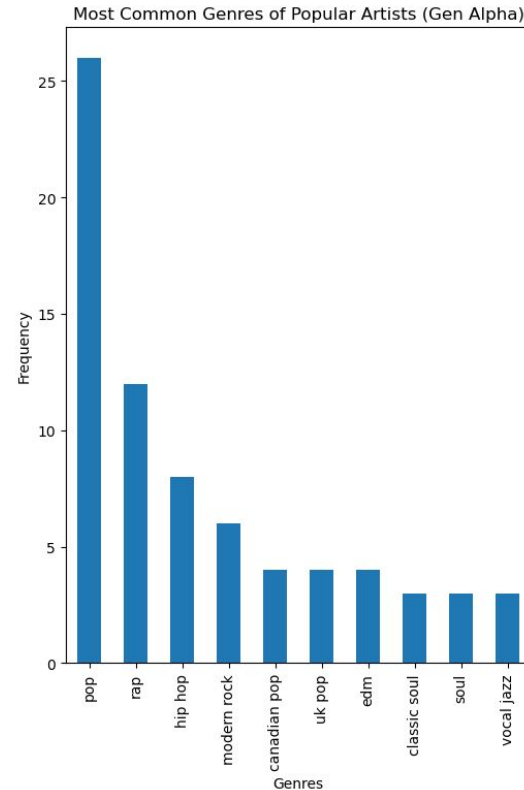
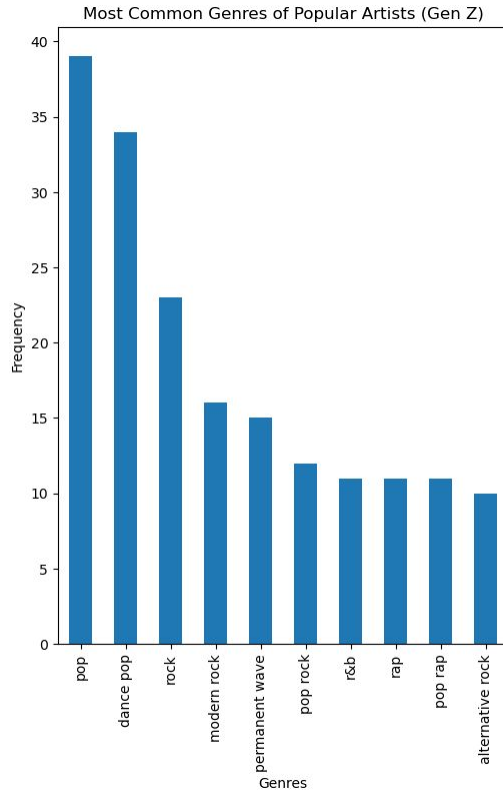
Most Common Genres of Popular Artists (Gen X)



Most Common Genres of Popular Artists (Millennials)

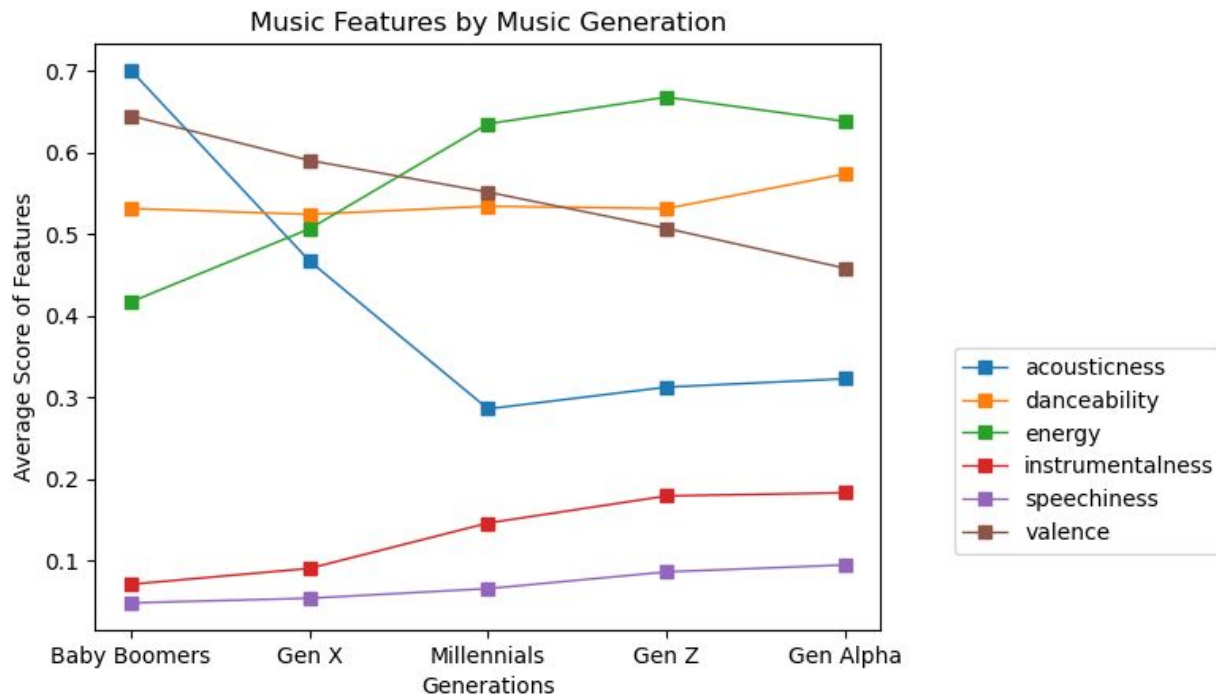


Top Artist Genres (cont.)

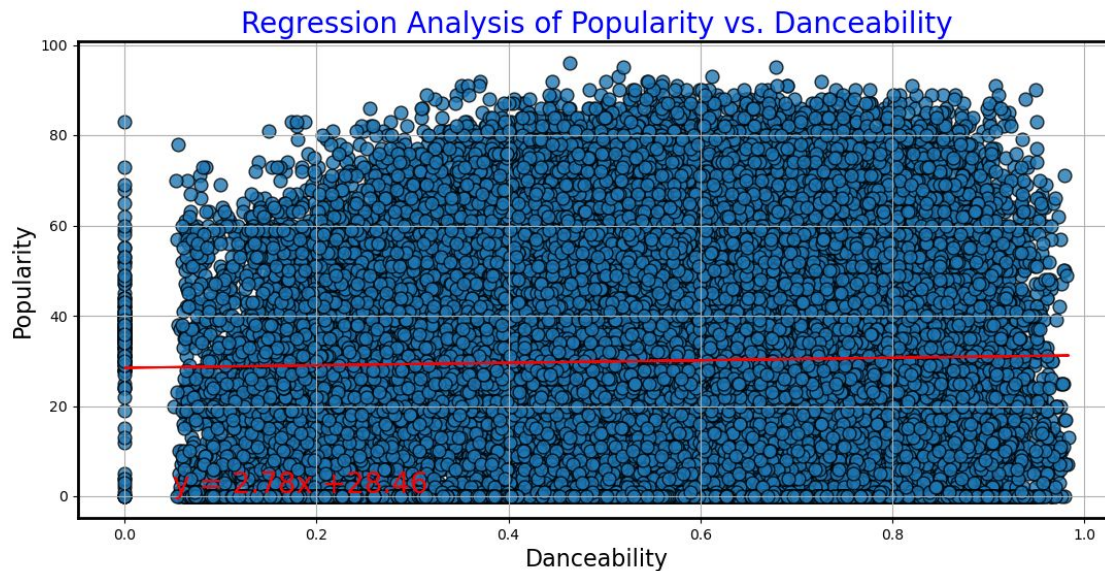


Music Features across Generations

Because we want to understand and potentially predict what music will sound like in the next generation, we look at what types of features of tracks are popular now and in the past.



Hypothesis and Testing



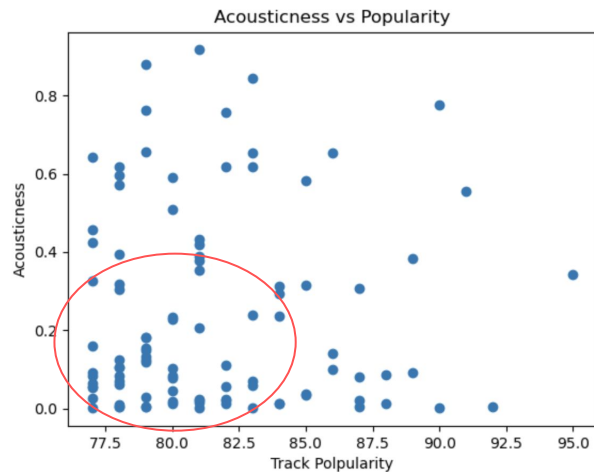
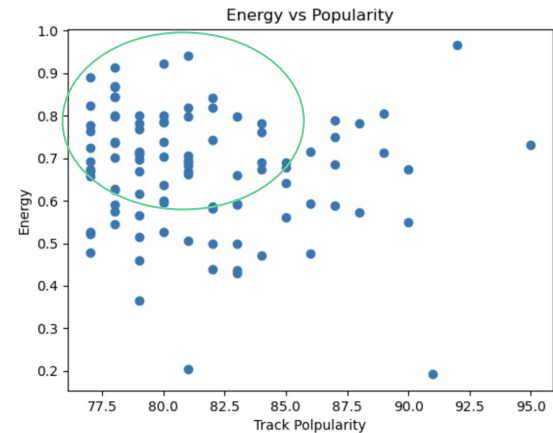
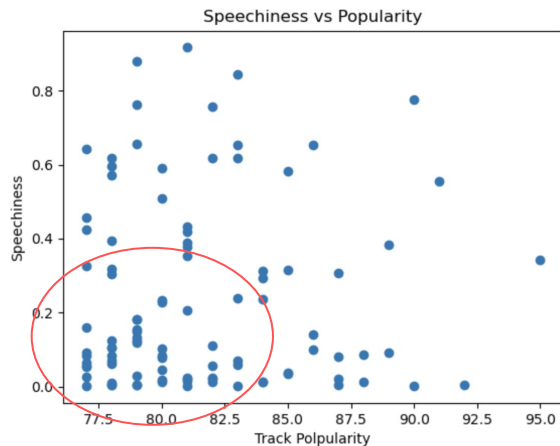
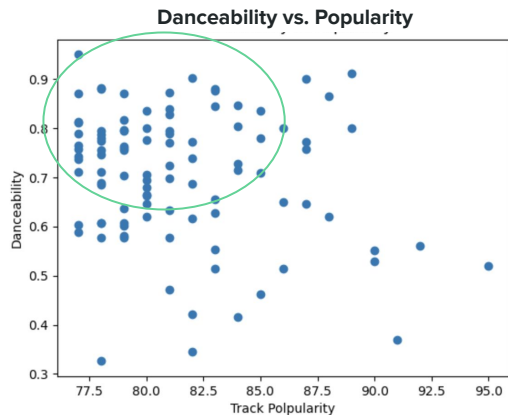
The **Correlation Coefficient** is: 0.02

The **r-value** is:
0.00051

The **p-value** is:
(3.4233363449656103e-1
0, 5)

We also look to see how one of the high rated features, danceability, correlates to a track's popularity. Based on the results of the regression, we see that one feature of a track alone does not necessarily warrant a popular song. Thus thinking that what makes a song popular is a combination of a song's features, the time in which it was released.

Hypothesis and Testing Cont'd



When we look at the Top 100 songs streamed each year since 2022, we see the top features across generations (Danceability, Energy, Valence) are clustered similarly with high scores of the features all having similar popularity scores (between ~77-85). Also aligning with the less favorable features like Speechiness and Acousticness having low scores when attributed to a song with a favorable popularity score.

Conclusion

Based on the trends of popular genres that transcend generations, as well as the measurement of various track features in each generation we believe we can predict the common trends for the next generation of music listeners

Prediction for Music of Generation Beta

Genre(s)	Common Track features	Variety of Music
Tracks are likely to be genre-bending/genre agnostic	Positive tracks that are able to be danced to with high energy	High volume of tracks to consume

APPENDIX

Audio Features**

Column Description

track_id: The Spotify ID for the track

artists: The artists' names who performed the track. If there is more than one artist, they are separated by a ;

album_name: The album name in which the track appears

track_name: Name of the track

popularity: The popularity of a track is a value between 0 and 100, with 100 being the most popular. The popularity is calculated by algorithm and is based, in the most part, on the total number of plays the track has had and how recent those plays are. Generally speaking, songs that are being played a lot now will have a higher popularity than songs that were played a lot in the past. Duplicate tracks (e.g. the same track from a single and an album) are rated independently. Artist and album popularity is derived mathematically from track popularity.

duration_ms: The track length in milliseconds

explicit: Whether or not the track has explicit lyrics (true = yes it does; false = no it does not OR unknown)

danceability: Danceability describes how suitable a track is for dancing based on a combination of musical elements including tempo, rhythm stability, beat strength, and overall regularity. A value of 0.0 is least danceable and 1.0 is most danceable

energy: Energy is a measure from 0.0 to 1.0 and represents a perceptual measure of intensity and activity. Typically, energetic tracks feel fast, loud, and noisy. For example, death metal has high energy, while a Bach prelude scores low on the scale

key: The key the track is in. Integers map to pitches using standard Pitch Class notation. E.g. 0 = C, 1 = C# / D b, 2 = D, and so on. If no key was detected, the value is -1

loudness: The overall loudness of a track in decibels (dB)

mode: Mode indicates the modality (major or minor) of a track, the type of scale from which its melodic content is derived. Major is represented by 1 and minor is 0

speechiness: Speechiness detects the presence of spoken words in a track. The more exclusively speech-like the recording (e.g. talk show, audio book, poetry), the closer to 1.0 the attribute value. Values above 0.66 describe tracks that are probably made entirely of spoken words. Values between 0.33 and 0.66 describe tracks that may contain both music and speech, either in sections or layered, including such cases as rap music. Values below 0.33 most likely represent music and other non-speech-like tracks

acousticness: A confidence measure from 0.0 to 1.0 of whether the track is acoustic. 1.0 represents high confidence the track is acoustic

instrumentalness: Predicts whether a track contains no vocals. "Ooh" and "aah" sounds are treated as instrumental in this context. Rap or spoken word tracks are clearly "vocal". The closer the instrumentalness value is to 1.0, the greater likelihood the track contains no vocal content

liveness: Detects the presence of an audience in the recording. Higher liveness values represent an increased probability that the track was performed live. A value above 0.8 provides strong likelihood that the track is live

valence: A measure from 0.0 to 1.0 describing the musical positiveness conveyed by a track. Tracks with high valence sound more positive (e.g. happy, cheerful, euphoric), while tracks with low valence sound more negative (e.g. sad, depressed, angry)

tempo: The overall estimated tempo of a track in beats per minute (BPM). In musical terminology, tempo is the speed or pace of a given piece and derives directly from the average beat duration

time_signature: An estimated time signature. The time signature (meter) is a notational convention to specify how many beats are in each bar (or measure). The time signature ranges from 3 to 7 indicating time signatures of 3/4, to 7/4.

track_genre: The genre in which the track belongs